A Section 508–conformant HTML version of this article
is available at https://doi.org/10.1289/EHP11438.

Response to "Comment on 'Deep Ensemble Machine Learning Framework for the Estimation of $PM_{2.5}$ Concentrations'"

Wenhua Yu,[1] Shanshan Li,[1] Tingting Ye,[1] Rongbin Xu,[1] Jiangning Song,[2] and Yuming Guo[1]

[1]Climate, Air Quality Research Unit, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[2]Monash Biomedicine Discovery Institute, Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Australia

https://doi.org/10.1289/EHP11438

Refers to https://doi.org/10.1289/EHP11385

We appreciate the comments by Stafoggia et al.[1] on our article.[2] We would like to clarify that our article aimed to introduce a novel methodology, deep ensemble machine learning (DEML), to estimate concentrations of air pollutants (e.g., daily, monthly, or annual average concentrations of fine particles, ozone, or nitrogen dioxide). We used Italian data for particulate matter ≤ 2.5 μg in aerodynamic diameter ($PM_{2.5}$) as a practice case to demonstrate how to implement DEML. The estimated concentrations of $PM_{2.5}$ will not be used for epidemiological studies.

We agree that the number of monitoring stations measuring $PM_{2.5}$ in Italy is larger than what we used. One major reason is that, as we stated, we deleted all missing values in the input predictors to reduce the uncertainties involved in interpolation. We included 204 monitoring stations in our original data set; after deleting the missing data, we were left with 133 valid monitoring stations. In addition, we acknowledge that predictions in southern Italy (including the islands of Sicily and Sardinia) may have high uncertainties because of fewer monitoring sites and high spatial variability, which we discussed in our paper as one of the limitations of our analysis.

In general, we agree that model performances, strictly speaking, should be compared independently with the same training and testing data sets, as we did in our analyses. It is better not to compare model performance with different training and testing data sets used in other studies. However, using even a relatively small number of monitors demonstrates the advantages and outstanding performance of DEML, compared with other commonly used single machine learning models, such as gradient boosting machine, support vector machine, random forest, and eXtreme gradient boosting.

As for the lack of some key spatial or spatiotemporal predictors, we agree that those predictors may be helpful for $PM_{2.5}$ estimations. However, our DEML model, with fewer predictors, still achieved high accuracy (coefficient of determination $= 0.87$). From another perspective, this reflects the ability of DEML to capture the complex relationships and spatiotemporal variations between $PM_{2.5}$ and predictors. We believe that, given more monitoring stations and other key predictors, the performance of DEML would be even better.

References

1. Stafoggia M, Cattani G, Ancona C, Gasparrini A, Ranzi A. 2022. Comment on "deep ensemble machine learning framework for the estimation of $PM_{2.5}$ concentrations." Environ Health Perspect 130(6):068001, https://doi.org/10.1289/EHP11385.
2. Yu W, Li S, Ye T, Xu R, Song J, Guo Y. 2022. Deep ensemble machine learning framework for the estimation of $PM_{2.5}$ concentrations. Environ Health Perspect 130(3):37004, PMID: 35254864, https://doi.org/10.1289/EHP9752.

Address correspondence to Yuming Guo, School of Public Health and Preventive Medicine, Monash University, Level 2, 553 St. Kilda Rd., Melbourne, VIC 3004, Australia. Email: yuming.guo@monash.edu. Or, Shanshan Li, School of Public Health and Preventive Medicine, Monash University, Level 2, 553 St. Kilda Rd., Melbourne, VIC 3004, Australia. Email: shanshan.li@monash.edu

The authors declare they have no actual or potential competing financial interests.